Abstract

Currently, the majority of substances tested in lifetime bioassays in rodents are not mutagenic and, therefore, at the most weakly carcinogenic, generally by epigenetic mechanisms. It thus appears obvious that only marginal increases of tumour incidences can be expected in lifetime bioassays and that, therefore, every aspect of a potential carcinogenic effect must be thoroughly evaluated. This paper describes a series of key factors, which should be looked at in order to exclude that the lifetime bioassay in question is flawed for design, technical or qualification reasons. It also provides some hints whether there is indeed a real effect and not just a variation of the spontaneous tumour incidences. Tumour findings must be seen in the context of the animal model, the pharmacokinetics and pharmacodynamics of the test substance, as well as any other observation in the present or other studies with the test substance, including non-tumour findings and—in particular—potential precursor lesions and effects on feed intake and survival. The possibility that the observed carcinogenic effects may be species-specific and not relevant for man is discussed. It is also important to check what findings are reported with similar substances or substances with the same pharmacological effect. Data from additional investigations on material of the same study and/or mechanistic studies are often needed to support the final risk assessment. © 2002 Elsevier Science Ireland Ltd. All rights reserved.

Keywords: Lifetime rodent bioassay; Carcinogenicity testing; Mechanism of tumour formation; Confounding factors; Risk assessment

1. Introduction

There are two basic types of tumorigenesis namely genetic and epigenetic (Weisburger and Wynder, 1984; Barrett, 1992; Williams, 1992; Couch, 1996; Melnick et al., 1996; Williams et al., 1996; Klaunig et al., 2000; Silva Lima and Van der Laan, 2000). Overtly genotoxic drugs are recognised early in development and, with the exception of certain classes, such as antineoplastic drugs, they are not further developed. Drugs,
which tested negatively in short-term mutagenicity and genotoxicity assays, may still have a weak carcinogenic potential in lifetime rodent bioassays, usually due to epigenetic mechanisms. The variability of spontaneous rodent tumour incidences (Mohr, 1999) and the large number of statistical tests performed on tumour data from lifetime rodent bioassays often result in equivocal results (Lin and Rahman, 1998). If small increases of tumour incidences achieve statistical significance, then one needs to differentiate between a biologically non-relevant finding and a true carcinogenic effect. This is particularly difficult in the case of a rare tumour or with a slightly increased incidence of a common tumour in high dose groups. Also, the possibility of a bell-shaped dose-response curve with biologically significant increases of tumour incidences in low or mid, but not in high dose groups cannot be overlooked (Calabrese and Baldwin, 1998; Teeguarden et al., 1998). If there are indications of a potential carcinogenic effect in a lifetime rodent bioassay, a number of questions need to be answered before one can provide a realistic assessment of human safety, i.e.:

- Is the study valid?
- Is there indeed an adverse effect?
- Are there other relevant findings?
- What is the (pharmacological) action of the compound in the test species and could the adverse effect be species-specific? Which organ is affected and is the organ relevant for man?
- What is known about other substances of the same class?
- Are additional investigations needed to establish the mechanism of tumour formation in the test system?

This article provides some details along these lines as well as guidance and hints to finding the answers to the questions formulated above. Some issues raised in this paper relate to the quality of lifetime rodent bioassays in general. Essentially the same process may be followed for the assessment of findings in carcinogenicity tests using transgenic mice, such as, e.g. p53 +/−, TgHras2 and XPA transgenic mice, or neonatal mice; however, it is important to take into account the special nature of these novel models (Harvey et al., 1993; Tennant, 1998; Tennant et al., 1995, 1999; Gollapudi et al., 1998; Rudmann and Durham, 1999).

The main focus of this review is on pharmaceutical preparations, but similar procedures can be followed for agrochemicals and other environmental agents. The paper is based on the expertise of the authors, supplemented by extensive reviews of the literature, and derived from lectures given to toxicologists at the University of Zurich and elsewhere. The paper is aimed at toxicologists and pathologists, who conduct lifetime rodent bioassays, analyse data from such studies or are involved in the reviewing process. The extensive literature base at the end of the paper is to ease access to further information if issues have arisen in lifetime rodent bioassays.

2. Is the study valid?

It is understood that this question needs to be asked for every study irrespective of its result. Several facets need to be considered.

2.1. Was the design and conduct of the study adequate?

In particular, is the species tested representative for man with regard to anatomy and physiology, absorption, distribution, metabolism and excretion (ADME) of the test compound, or to the pharmacodynamic action of drugs? The use of albino rodents is not appropriate, if toxicity studies in pigmented animals (dogs, minipigs, etc.) have shown a particular affinity of the test compound to the pigmented cells, especially if associated with pathological changes in shorter term studies. Is the breeder, who supplied the test animals, reliable (Engelhardt et al., 1993)? Was the quality of animals adequate with regard to latent infections (Maurer et al., 1993)? Was the husbandry good and the diet nutritionally sufficient (Zusman, 1998) and without contamination?

One key factor is dose selection (Butterworth et al., 1995; Foran, 1997; Food and Drug Administration, 1995, 1997).
the dose scientifically sound (Apostolou and Helton, 1993; Monro and Davies, 1993; Haseman and Lockhart, 1994; Bergman et al., 1998)? What are the credentials of the study director (Ettlin and Hodel, 2000; Ettlin et al., 2001)?

2.2. **Was the test substance analytically characterised according to current standards?**

How pure was it and what were the impurities (e.g. Food and Drug Administration, 2001)? Could an impurity (by-product, degradation product or contaminant) have been responsible for the tumour response seen?

2.3. **Were the technical pathology procedures appropriate?**

This includes consistent necropsy, sampling and histotechnical procedures (Bahnemann et al., 1995). Standardised processing is essential. Tissue accountability is an important factor to judge the quality and validity of a lifetime bioassay. It is influenced by various factors, such as autolysis, age of the animal (e.g. thymus), size of organ, amount of tissue available and preserved during necropsy and embedded in paraffin blocks. Although there are no guidelines, a reasonable minimum standard is that 90% of protocol organs/sex/group should be examined microscopically to ensure study validity.

2.4. **Were the diagnostic pathology procedures appropriate?**

The most important lesions in lifetime rodent bioassays are tissue proliferations: hyperplasia, benign tumours and malignant tumours. The biological behaviour of these lesions is often not well known, as—in contrast to human pathology—no biopsies and follow-ups are available. This may also render the establishment of the cause of death difficult (Ettlin et al., 1994; Haseman, 1995).

Histopathological diagnosis of proliferative lesions in rodents is partly based on assumptions and agreement among pathologists, including size criteria to label a tumour as benign or malignant (Ettlin et al., 1992). Therefore, standardisation of diagnostic criteria is indispensable and the method should be referenced in the report (Mohr et al., 1990; Mohr and Morawietz, 1995). Deviation from standardised diagnostic criteria is permissible, but must be justified and defined. Examples of standardised nomenclature systems are: WHO/IARC/RITA nomenclature: the rat (Mohr et al., 1992–1997); the mouse (Mohr, 2001); STP nomenclature (STP/ARP/AFIP, 1990–1998); IARC nomenclature for rats (Turusov and Mohr, 1990) and mice (Turusov and Mohr, 1994); monographs on the pathology of laboratory animals (Jones et al., 1983–1996).

The diagnostic quality depends on the pathologist, who is generally expected to have an academic degree in veterinary or human medicine. To be qualified to read long-term bioassays he/she also needs many years of postgraduate training in pathology. Board certification (US, EU, International Academy of Toxicologic Pathology IATP) or registration for toxicologic pathology (EU, J) is another hallmark of qualification in pathology (Ettlin and Hodel, 2000; Ettlin et al., 2001). Generally, one pathologist should read all histological slides of a study. If for deadline reasons two pathologists have to share the task, then splitting by males and females may be considered, while splitting by dose groups should be avoided. All slides from one animal are generally read at the same time, as the various lesions are interdependent and often have biologically significant effects on the overall condition of the animal.

Blinded slide evaluation (i.e. without knowing the treatment group) is not recommended for routine examinations (STP, 1986). However, blinded reading of an organ with an equivocal or subtle finding can be very helpful. If done, it should be recorded in the raw data and the study report and this will enhance the confidence of the regulatory reviewer on the validity of the data.

The recording method for tumour multiplicity and for concomitant occurrence of hyperplasia and neoplasia in the same organ or system must be stated in the method section. No general recommendation can be given.
Should grading of tumorigenicity be attempted? Possible criteria are the number of tumours/animals with tumours, or the size and malignancy of tumours. Practical importance of such grading is generally small.

2.5. Guidelines

Guidelines regarding the design and conduct of lifetime rodent bioassays were harmonised essentially world-wide (e.g. for drugs by various International Conferences on Harmonisation) and are available from the authorities (e.g. Food and Drug Administration, 1998; OECD, 2001) or from scientists working for government institutions (e.g. Jarabek and Farland, 1990; Lai et al., 1994). Guidelines were also published by some scientific and professional societies such as the International Federation of Societies of Toxicologic Pathology (Faccini et al., 1992), by the Drug Information Association (Spindler et al., 2000), the Centre for Medicines Research (McAuslane et al., 1992), groups of toxicologists and pathologists (RITA, 1999) and individuals (Monro and MacDonald, 1998).

Strict adherence to GLP as requested in the regulatory guidelines is essential. Overall, the requirements regarding expertise and infrastructure are so complex that only a small number of facilities is able to conduct lifetime rodent bioassays.

3. Is there indeed an adverse effect?

It is a good practice when an experienced pathologist reviews a selection of histopathological slides from lifetime rodent bioassays (STP, 1991; Ward et al., 1995; Long, 1996). The extent of such a peer review depends on the study. It is recommended to review at least 10% of all tumours and all target organs. However, some companies review all tumour findings. In addition, a complete review of 10% of animals/group serves to ascertain that no significant lesions were missed. The peer review needs to be documented in raw data and study report.

3.1. Is the analysis of tumour incidences scientifically sound?

The analysis of tumour findings includes a correlation of the available pathology findings with relevant in-life data and in particular the first occurrence of palpable masses. It is also a good practice to establish the cause of death for animals that died or were killed in moribund state, even if the statistical test used for data analysis does not require cause of death data (Ettlin et al., 1994). Increased and decreased tumour incidences should be recorded and discussed in relation to treatment (Davies and Monro, 1994; Haseman and Johnson, 1996; Linkov et al., 1998a,b).

There are various approaches to evaluate the tumour incidences. Incidences can be analysed per dose group in total (males and females combined) or per dose group and sex (males and females separate). If two control groups were used, statistical pair-wise comparisons with the treatment groups can be done with pooled and single control groups. Decedent and terminal sacrifice animals can be combined or analysed separately. Note that groups can only be combined if there is no reason to assume that they represent different populations, which needs to be shown by statistical testing (Vater et al., 1993). Summary figures, such as the number of animals/group with tumours in total or with single or multiple tumours, as well as the total number of tumours/group are often not as helpful as are more detailed figures (Haseman et al., 1986). Time to appearance of tumours is another important parameter, but generally only assessable for externally palpable tumours (Dankovic et al., 1993). An effort should also be undertaken to examine the effect of induced tumours on the life span of the test animals (Ettlin et al., 1994).

Statistics play an important role in the evaluation of tumour incidences. The following tests are routinely used:

- Fischer test: pair-wise comparison; good for small numbers of tumours or animals
- Peto test: survival-adjusted trend test (asymptotic form) stratified according to time (often in four periods); need to distinguish fatal and non-fatal tumours; needs larger number of animals (50+)
Exact Peto: combination of Peto and Fischer test; good for small numbers of tumours or animals
Cochran–Armitage test: trend test (simpler than Peto test); no need to determine cause of death; often used for non-tumorous lesions


To increase the statistical power of rare observations, often one has to combine the incidences of similar tumours, e.g. tumours with similar histogenesis such as benign and malignant tumours of the same cell type in the same organ/organ system (e.g. liver adenoma and carcinoma). At times it is reasonable to combine tumours also with precursor lesions, particularly when the distinction between hyperplasia and tumour is based essentially arbitrarily on the size of the lesion, as e.g. for interstitial cell hyperplasia and benign tumour.

The following proliferative lesions can generally be combined (McConnell et al., 1986):

- neoplasms in the respiratory tract or in the peripheral lung, but generally not from both locations together;
- squamous cell neoplasia of the upper alimentary tract including forestomach;
- epithelial neoplasm of the small intestine and the large intestine and—depending on the type of proliferative lesions—also of small and large intestine;
- smooth muscle neoplasia body-wide with exception of those of the GI tract and reproductive tract, which must be evaluated separately.

The above rules depend partly on the strain of animals used. What type of proliferative lesions can be combined has often to be determined on a case-by-case basis and keeps changing with the evolving understanding of the mechanism of tumour formation.

3.2. What is the historical tumour incidence?

A rodent lifetime bioassay is an epidemiological study and the concurrent control is the best control. Sometimes the concurrent control group may show an unusual tumour incidence (low or high) or rare tumours, not present in the concurrent control group, may be found in one of the dose groups. Under such circumstances, one might use historical control data from the same strain and facility. Ideally, these data should not be older than 5–8 years, because of the genetic drift and evolution of diagnostic nomenclature. If sufficient data are not available from the same facility, one might consider using data from a historical control database (Huff and Haseman, 1991; McAuslane et al., 1993). Good historical control databases are characterised by the following (Roe, 1994; Haseman, 1995; Mohr, 1999):

- quality control (consistent nomenclature, peer review, data check);
- data on possibly modifying factors, such as strain, breeder, husbandry, age of the animals, sampling and histotechnical procedures;
- audit trail for changes of nomenclature over time available;
- data on stability of tumour pattern over time.

Visual documentation of characteristic lesions in the database are a special asset (Morawietz et al., 1992).

4. Are there other relevant findings?

There is a number of factors, which can modify the outcome of a lifetime rodent bioassay (Hardisty, 1985; Munro et al., 1995; RITA, 1999; Calabrese, 2000a), the most important being feeding. Food intake must be analysed and taken into consideration when evaluating lifetime rodent bioassays.

- It is known that feeding ad libitum significantly increases tumour incidences of the pituitary, the mammary gland and the lung in rats and of the liver in mice (Salmon et al., 1990; Roe et al., 1995; Seilkop, 1995; Klaasen, 1999). However, there is a report of decreased uterine tumours in rats (Roe et al., 1995). Feeding ad libitum also significantly increases the incidence and severity of degenerative diseases, including nephropathy, myocarditis, polyarteritis and prostatitis (Turnbull et al., 1985). Nephropathy might significantly impact on the excretion of the test compound. Overall, feeding ad libitum shortens the life span of test animals.
On the other hand, reduced feeding (Keenan et al., 1995; Christian et al., 1998; Keenan et al., 1998) or feeding ad libitum of modified diets (Rao, 1996) significantly decreases incidences of most proliferative lesions, in particular those of endocrine organs. It also lowers the occurrence and severity of age-related degenerative diseases. However, incidences of tumours of testis and uterus in rats may be increased (Roe et al., 1995). Caloric restriction can decrease the metabolism of xenobiotics and herewith modify the outcome of lifetime bioassays (Manjgaladze et al., 1993).

Unintended, but not infrequent in lifetime bioassays (Gaylor and Kodell, 1999), dosed animals eat less than control animals because of toxicity and/or bad palatability of the test compound given by feed admixture.

Carcinogenicity data must also be evaluated against the background of toxicity data from the same and other studies with the same compound. Are there any preneoplastic lesions? Slides of earlier studies can be re-examined, as, with hindsight, lesions may be detected which were originally not considered significant. Precursor lesions, which have the potential to give rise to tumours, include:

- Adaptive changes resulting from chronic tissue irritation. An underlying reason can be crystallised test substance forming aggregates with proteins and lipids from body fluids in the excretory system; e.g. saccharine fed for lifetime to male rats results in urinary bladder carcinoma (Cohen and Lawson, 1995; Whysner and Williams, 1996a). Solid state carcinogenicity in rats at subcutaneous injection sites (Grasso et al., 1991) or in mice at implantation sites of microchips for animal identification (Tillmann et al., 1997) are other examples of the consequences of chronic tissue irritation in rodents, which are of little relevance in man.

- Hypertrophy and hyperplasia. An increase in cell size and an increase in cell number are not always easy to distinguish without special investigations and may in fact often occur in combination. Cells can become hypertrophic because of cell organelle proliferation, such as proliferation of peroxisomes seen with hypolipemic drugs (e.g. clofibricates), which are associated with liver tumours in rats but not man (e.g. Ashby et al., 1994; Kluwe, 1994; Lake, 1995; Melnick et al., 1996; Richert et al., 1996; Cattley et al., 1998). Proliferation of the smooth endoplasmic reticulum is often the reason for cell hypertrophy (Garg et al., 1979; Kanai et al., 1993) and is often associated with liver tumours in rodents but not in man (Butterworth et al., 1995; Whysner et al., 1996). Liver enzyme inducers can also lead to thyroid tumours because of increased T3 metabolism and the consequent TSH increase (McClain, 1992, 1995; MacDonald et al., 1994). Again, this is of no relevance to man. Further examples of hyperplasia following hormonal and other receptor-mediated stimulation are given below under ‘What is known about other substances of the same class’.

- Cytotoxicity and necrosis may lead to cell proliferation as an expression of tissue repair. For example, various substances bind to α2μ-globulin in renal tubules, thereby leading to kidney toxicity and tumours in rats (Swenberg, 1993; Hard and Whysner, 1994; Whysner and Williams, 1996b).

Irrespective of the underlying pathogenesis, cell proliferation observed in shorter term studies at comparable doses generally indicates an increased likelihood for finding tumours in lifetime rodent bioassays (Melnick et al., 1993, 1996; Huff, 1995; Cohen, 1995a). However, cell proliferation is not always followed by neoplastic changes (Ward et al., 1993).

Other carcinogenicity data should also be considered, e.g. data from a conventional carcinogenicity study in a second species, and eventually from a carcinogenicity study in genetically modified or neonatal mice or from other short-term carcinogenicity studies. Additional indications for tumorigenicity of a substance may be available from quantitative structure activity relationship data (QSAR, Richard, 1998). Results of genotoxicity testing (Ashby, 1991, 1992; Rosenkranz and Klopman, 1993; Benigni et al., 1996) may need to
be re-evaluated carefully in the light of tumour findings in lifetime rodent bioassays. The majority of genotoxic compounds is also carcinogenic (Zeiger, 2001). However, most animal carcinogens tested in lifetime rodent bioassays act by epigenetic mechanisms and are not genotoxic, particularly those producing tumours in rats only (Table 1). Equivocal short-term genotoxicity in the presence of equivocal lifetime bioassay data weighs relatively heavy in safety assessment.

5. What is the (pharmacological) action of the test compound in the test species and could the adverse effect be species-specific?

5.1. Rodents are not men

Physiology of rodents is in several respects quite different from man (Grisham, 1997). Particularly, the endocrine regulation of rodents is different in many aspects and easily disturbed. Examples include: rats lack high-affinity thyroxin-binding globulin (Hamolsky, 1969; Alison et al., 1994); the oestrogen/progesterone ratio is 1:100 (~200) in rats, but 1:1 in women; reproductive senescence means progesterone dominance in old rats, but a waning system in women (Richardson et al., 1984); prolactin has a trophic effect on the mammary gland of rats, while it maintains lactation in women (Neumann, 1991); rat, but not human Leydig cells react with tumour formation following long-term stimulation by high LH levels (Prentice and Meikle, 1995).

Table 1
Concordance between bioassay and genotoxicity findings based on data published in the PDR (Davies and Monro, 1995)

<table>
<thead>
<tr>
<th>Bioassay</th>
<th>Total number</th>
<th>Genotoxic (% of total number)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive in both rats and mice</td>
<td>39</td>
<td>10 (26%)</td>
</tr>
<tr>
<td>Positive in rats only</td>
<td>37</td>
<td>2 (5%)</td>
</tr>
<tr>
<td>Positive in mice only</td>
<td>22</td>
<td>4 (18%)</td>
</tr>
<tr>
<td>Negative in both rats and mice</td>
<td>123</td>
<td>1 (0.8%)</td>
</tr>
</tbody>
</table>

Further differences between rodents and man may be based on pharmacokinetic differences (Green, 1991; MacDonald et al., 1994; Lake, 1999). Therefore, among other aspects the following ones need clarification: which metabolites are formed and could they be genotoxic (see e.g. tamoxifen, Swenberg, 1997)? Does the compound have a special affinity for a particular organ? How is the compound excreted?

The influence of the species is also reflected in the wide variation of incidences of spontaneous pathology findings among different species. Mice have a high tumour incidence in lung and liver and partly in Harderian and adrenal glands, the hematopoietic system and ovaries (Sher et al., 1982; Lohrke et al., 1984; Melchionne et al., 1986; Maita et al., 1988; Tamano et al., 1988; Bombhard and Mohr, 1989; Rittinghausen et al., 1997; Tillmann et al., 2000). Rats have a high incidence of mammary gland and pituitary tumours (Sher et al., 1982; Bode et al., 1985; Poteracki and Walsh, 1998). Rats and mice also differ partly in their response to lifetime exposure to chemical substances (see Table 2), but in the majority of cases data of carcinogenicity data from mice and rat lifetime bioassays are in agreement (Young, 1989).

There are also marked differences in tumour incidences between different strains of the same species, in particular for:

- Leydig cell tumours (88–96% in F344 rats, below 10% in SD rats, 1–2% in Long–Evans rats; Prentice and Meikle, 1995).
- Mononuclear cell leukaemia (20% in F344 rats, rare in SD rats; Abbott et al., 1983).
- Mammary gland tumours, the incidence of which varies widely between different rat strains, possibly due to endocrine differences (Sinha et al., 1975) and viral infections (Nandi et al., 1966).
Significant intra-strain differences exist also between animals from different breeders (Engelhardt et al., 1993) and even between animals from the same breeder (Wolff, 1996).

5.2. Which organ is affected and is the organ relevant for man?

Species, strain, breeder, dose and route of application of the test substance influence the organ-specific tumour incidence (Morawietz and Rittinghausen, 1992). Overall there are no marked differences in the distribution of tumours induced by mutagens and non-mutagens (Ashby and Patton, 1993; Gold et al., 1993), though differences were described for some substances, such as chloromethine (Bucci, 1985; Brock et al., 1989). Epigenetic carcinogens usually increase tumour incidences in organs with a high spontaneous rate of tumours and lead to neoplasms with similar histology as their spontaneous counterparts (Hardisty, 1985). This is further supported by data published in the Physician’s Desk Reference (PDR; Davies and Monro, 1995), i.e. the most frequent target organs were:

- in rats: thyroid, liver, testis, mammary gland, adrenal and pituitary;
- in mice: liver, lung, mammary gland, bone marrow and ovary.

Small but statistically significant increases in the above tumour incidences are often regarded as less relevant than the appearance of new tumour types. This is the case with tumours in livers (Roe, 1987), particularly in mice livers (Alden et al., 1996; Ván Oosterhout et al., 1997; Ekman, 1998; Liteplo et al., 1998).

Rodents have additional anatomical particularities. The following organs are not found in man: forestomach, Harderian gland (eye region), Zymbal’s gland (ear) and preputial gland. Tumours occurring exclusively in these organs are often regarded as not relevant for man (Whysner and Williams, 1996c), although tumours in the forestomach might indicate, for example, a risk for oesophageal tumours in man and the similarity of Zymbal’s and preputial glands to human sebaceous glands must be kept in mind.

6. What is known about other substances of the same class?

Literature searches and data from regulatory agencies obtained under the Freedom of Information Act may be useful to interpret findings in a lifetime rodent bioassay. There are a number of drugs in the market, which are weakly carcinogenic in animals and humans because of their mode of action (e.g. alkylating anticancer drugs; IARC, 2000) and pharmacological effect (e.g. immunosuppressive drugs; Pavelic and Hrsak, 1978; Yabu et al., 1991; Gruber et al., 1994; Penn, 2000). Several classes of drugs are known to be animal carcinogens but are regarded as safe for man (Heywood, 1987; Davies and Monro, 1995). Some have been mentioned above, when discussing precursor lesions (see Section 4) and include peroxisome proliferaters such as hypolipemic drugs (clofibrates) and compounds inducing drug metabolising enzymes (best known representative: phenobarbitone). Other examples are:

- Hormones and other compounds with some endocrine (side) effects: hyperplastic and neoplastic changes are often seen following hormonal and other receptor-mediated stimulation (Roe, 1989). Bromocriptine and mesulergine, both dopaminergic compounds that lower prolactin levels, significantly decrease the incidence of endocrine tumours but increase that of uterine tumours (bromocriptine: Richardson et al., 1984) and of Leydig cell tumours (mesulergine: Prentice et al., 1992). It is also known that tumours of endocrine organs are often associated with tumours of the adenohypophysis.

- Compounds, which—by their intended pharmacological action—block the normal function of an endocrine organ or system, often lead to tumours in this organ or system: e.g. H2-blockers lead to ‘carcinoids’ in the rat stomach (Hirth et al., 1988) and antithyroid agents result in thyroid tumours (McClain, 1992, 1995). These tumours are without relevance to man.

- Similar findings are reported with other compounds, which disrupt endocrine function through an adverse action: e.g. sulphonamides
interfere with thyroid function and lead to thyroid tumours only in rodents (Swenberg et al., 1992; McClain, 1992, 1995).

- β2-agonists, such as salbutamol or terbutaline, lead to hyperplasia of salivary glands and mesovarian leiomyomas in rats, the pathogenesis of which is not well understood (Colbert et al., 1991). Their occurrence can be suppressed by simultaneous administration of a β-agonist, such as propranolol.

The list of animal and human carcinogens is regularly updated in the IARC monographs (see e.g. IARC, 1987, 2000; Torok et al., 2000).

7. Are additional investigations needed to establish the mechanism of tumour formation in the test system?

Knowledge of the mechanism of action helps considerably in the assessment of tumorigenic findings in lifetime rodent bioassays (Roe, 1988; Schou, 1990; Gold et al., 1992; MacDonald et al., 1994; Butterworth et al., 1995; McClellan, 1995; Williams, 1997; Food and Drug Administration, 1998). Additional investigations can sometimes be performed on samples already available from the study in question or from previously conducted toxicity studies with the same compound (Ettlin et al., 1991). Examples for possible investigations include: blood samples for hormones (Prentice et al., 1992); blood or tissue samples for gene or mRNA expression or for marker proteins or protein patterns by proteomics (Chibout et al., 2002); tissue samples for investigations by EM or by immunohistochemical methods, e.g. for identification of the proliferative cell type (Oberholzer et al., 1987; Qureshi et al., 1991), measurement of cell proliferation (by PCNA, Goldsworthy et al., 1991; Perentes et al., 1994) or apoptosis (Schulte-Hermann et al., 1999); and morphometry (Ettlin et al., 1993).

Sometimes it is necessary to run additional mechanistic studies. Types of in vivo studies include: reversibility studies, cell proliferation studies (Lardelli et al., 1994), or EM investigations of appropriately fixed tissue samples (Ettlin et al., 1991). Special in vitro studies include cell or tissue cultures, organ slices or the perfused target organ itself to investigate metabolism, effects on subcellular organelles or on gene expression, or dose-effect relationship at subcellular or molecular levels (Mehlman et al., 1989). Receptor studies (Silbergeld, 1994; Schwarz et al., 1995) can be helpful, while studies of intercellular communication (Yamasaki, 1996) and initiation-promotion assays (Tsuda et al., 1999) may not contribute much useful data. The design of mechanistic studies is important to obtain relevant answers.

8. Discussion and conclusion

There can be no doubt that lifetime rodent bioassays are not a completely reliable test system (Roe, 1998). These studies are difficult to validate and are sometimes questioned with regard to their usefulness (Monro, 1996; Slovic et al., 1997). Nevertheless, lifetime rodent bioassays and short-term carcinogenicity studies in genetically modified or neonatal animals are widely accepted (McAuslane et al., 1992; Huff, 1999). Low sensitivity to detect a carcinogenic effect (Haseeman and Lockhart, 1994) is of higher concern than low specificity, i.e. carcinogenic findings of no relevance to man (Lovell, 1989; Omenn et al., 1995). The latter can be addressed as discussed in this paper.

The old question of how to interpret high dose findings and of how to extrapolate animal findings to man, is still not finally answered (Brown et al., 1988; Lutz, 1990; Gold et al., 1992; Olin et al., 1997; Calabrese, 2000b). For genotoxic and mutagenic compounds extrapolation of animal data to man is generally assumed to be possible with no threshold, below which exposure is considered safe (Cohen, 1995b). Epigenetic tumorigenicity is often species-specific and is considered to have a threshold and therefore, many rodent carcinogens are considered safe for man (Castro et al., 1993; Cohen and Ellwein, 1993; Purchase and Auton, 1995; Lutz, 2000). For overall safety assessment, often only the weight of evidence approach is possible (Food and Drug Administration, 1998) and the relevance of tumour findings has to be evaluated on a case-by-case basis (Kuschner, 1995). One also notes a considerable evolvement.
over time regarding the assessment of tumour findings, which has become less schematic and more pathogenesis-oriented (Scheuplein, 1995; Velazques et al., 1996). On the other hand, much discussion is ongoing regarding the shape of dose-response curves. The possibility of e.g. bell-shaped dose–response curves cannot be discounted a priori (Calabrese and Baldwin, 1998; Teeguarden et al., 1998). Compounds have various effects, which could at least theoretically counteract each other (Calabrese, 2000b): e.g. a carcinogen might induce less tumours at higher doses because of toxicity; splenic toxicity is often associated with decreased incidences of leukaemia in mice (Haseman and Johnson, 1996).

If an epigenetic carcinogenic effect of possible relevance to man is found, the single most important parameter to evaluate is the No-Toxic (Carcinogenic) Effect Level (NTEL) in relation to the exposure. Is there a sufficient safety factor defined as NTEL in the most sensitive species (as far as this species is of relevance to man) and the maximal human exposure (dose/concentration and duration of exposure) occurring or intended?

The use of the substance, for drugs in particular their therapeutic indication, is very important. Lower safety factors are acceptable for life saving indications of drugs than for not live-saving indications. In general, higher safety factors are needed, when children or young adults are treated, compared with those needed for patients towards the end of their natural life span. Are potentially safer alternatives available or has the substance in question a true advantage over the alternatives that would justify a minimal risk? How did regulatory authorities assess similar findings with other drugs? For chemicals other than drugs, generally no or only few human data are available. Furthermore, exposure to them happens by accident or contamination. Therefore, risk assessment tends to be more conservative for agrochemicals (Rakitsky et al., 2000), other environmental chemicals (Butterworth et al., 1995; Fan et al., 1995; Olin et al., 1997), consumer products (Babich, 1998), residues in food (Hazelwood, 1987) and chemicals at the workplace (Forth, 1996).

Conduct, analysis and interpretation of lifetime rodent bioassays is among the most demanding tasks a toxicologist or pathologist is faced with (Clayson et al., 1983; Nordic Council on Medicines, 1987). At every stage, much can go wrong and much money (up to and exceeding US$ 500,000) and development time (up to and exceeding 2 years) can be lost. Worse, the development of a compound, with expenses in excess of 500 mio. US$ up to this stage, can be jeopardised just before introduction to the market. It is good to remember that there are many more animal carcinogens than human carcinogens (Davies and Monro, 1995; Van Deun et al., 1997). On the other hand, toxicologists and toxicologic pathologists have an obligation to make sure that humans and the environment are not exposed to real carcinogens.

Acknowledgements

Many thanks to Christa Fletcher for secretarial support, to Elias Perentes for reviewing and to Regula Ettlin for the final checking of the manuscript.

References


Schwarz, M., Buchmann, A., Stinchcombe, S., Luebeck, G., Moolgavkar, S., Bock, K.W., 1995. Role of receptors in


